The Proposal

The Simplified Mongolian Solution – Unicode
Simple, Stable, and Searchable

Prepared by “Z” Team
Chuck.Namujila@gmail.com
09 Mar 2020
# Executive Summary

| Objective | To apply separate Usgiin (ᠠᠪᠧᡙᠧᠪᠧᡉ Үсгийн 音素) Unicode Blocks as the secondary options for:  
|           | • Hudum Mongolian script (128 code points)  
|           | • Todo Mongolian script (128 code points) |
| IT Investment | Self-sponsored & Zero Dollar investment from Government |
| Project Value - IRR | Unlimited, national strategic project |
| Project Timeline | 1 ~ 2 years |
| Scope | Out of Scope:  
| | • The U1800 Phonetic Mongolian Solution  
| | In the Scope:  
| | • The New Usgiin (ᠠᠪᠧᡙᠧᠪᠧᠧ Үсгийн 音素) Solution  
| | • Hudum and Todo Mongolian scripts |
| User Community Support | 10 thousand users signed up for an alternative solution in 2019 |
1-1. Problem Statement

The phonetic solution violated at least 4 out of 10 rules from Unicode Standard Chapter 2.2

Architecture Design principles:

• Violates “Universality” principle and the scripts are unreadable in historic writing systems
• Violates “Unification” principle, creates Code-Duplication, and become un-searchable
• Violates “Plain Text” principle and depends on FVS/MVS/Grammar rich text logic
• Violates “Stability” principle and becomes un-stable between versions

Known Facts:

• U1800 Phonetic script documents written in Microsoft word 2010 are showing DIFFERENT characters in Microsoft word 2016.
1-2. Problem Statement – “MVS/FVS” and “NNBSP”

- Visible foreign characters like “MVS” greatly demoralizes the Mongolian Users, in Red
- Invisible characters like “NNBSP” creates a lot of wrongs words, in Orange

Source: [www.President.mn/MNG](http://www.President.mn/MNG) using iPhone
1-3. Problem Statement – Dependency on Un-unified Grammar

- There are disagreements on Grammar rules across regions and academic streams. Academic stream A supports “Ail”, while Academic stream B supports “Ayil”

Source: Liang JinBao 《蒙古文正字法与蒙古文编码》
1-4. Problem Statement – Cybersecurity Phishing Risk

- Phishing risk for using “Yin” to replace “Jin”
1-5. Problem Statement – Code Duplication

There are 6 sets of code duplication in the current Phonetic solution:

1. a/e/n
2. e/o
3. h/g
4. i/y/j
5. o/u
6. t/d

7168 different options to spell.
1-6. Problem Statement – OpenType Dependency

Python Anaconda Spyder is not supporting the Phonetic solution.
(Note: This is one of the most popular Artificial Intelligence, machine learning software)
2. Urgency: Time is running out – Decreasing User Number

The Mongolian language is classified as one of “Endangered language in China” by UNESCO.

- Statistics shows the Mongolian user number deceased from 3.2 million in 1994 to 1.7 million in 2016.
- By straight line forecasting, the user numbers will be ZERO by 2036 in China domain.

Average losing 188 users per day
3-1. Architecture Design

Mongolian language is a Phonetic language, but the script should not violate Unicode Standard Architecture Design principles to keep the Phonetic feature.

- The underlying Unicode script should follow the Unicode Standard.
- The IME (Input Method Editor) could enable the Phonetic features.
3-2. Architecture Design – Descoping

- Moving the Phonetic feature and Grammar/variable rules into IME (Input Method Editor).
- The new Usgiin Unicode encoding is plain text, simple/stable/searchable
4-1. Recommendation - The Simplified Hudum Mongolian Solution

Z-55 solution is proposed by Z-team.

- Deposited into the smallest component.
- No grammar dependency, No code-duplication, No OpenType dependency, No FVS/MVS
- IME Input method will carry the Grammar & Variable Rules

Objectives:
- Moving away the Grammar problems from Unicode
- Removing Code duplication
4-2. Recommendation - The Simplified Todo Mongolian Solution

Mr. Sainbilig proposed the S-105 simplified Todo solution

- Deposited into the smallest component.
- IME method will be developed on Microsoft Windows, Mac, iPhones, and Android.
4-3. Recommendation - The Simplified Manchu Solution

To be updated after Hudum Mongolian and Todo scripts are approved by UTC.
5. Other Information

The simplified solution is to reduce the pressure on Phonetic solution development, buying more time and protecting the user community, until Phonetic solution solves the existing problems.

Different blocks for different scripts are suggested: Hudum script, Todo Script, Manchu script, Xibo script are all different and suggest to encode in Separate blocks.

IME (Input Method Editor) methods will be developed in separate projects to support the Simplified solution. The development ownership is NOT on Unicode.

OpenType is Optional:
• Without OpenType support, the simplified solution will ensure the accuracy.
• With OpenType support, the simplified solution Glyph will be gorgeous.

Z team is willing to offer a database with 3 Million Mongolian word/word-group records.
6. Q&A
Thank you